

25.06.99

日 本 国 特 許 庁

PATENT OFFICE
JAPANESE GOVERNMENT

REC'D 13 AUG 1999

WIPO PCT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日

Date of Application:

1998年 6月 4日

出 願 番 号

Application Number:

平成10年特許願第155550号

出 願 人

Applicant(s):

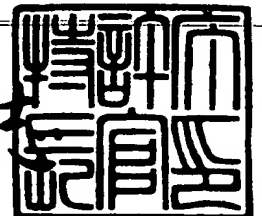
松下電器産業株式会社

PRIORITY
DOCUMENTSUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

1999年 7月15日

特許庁長官
Commissioner,
Patent Office

山 佐 平



出証番号 出証特平11-3050073

【書類名】 特許願

【整理番号】 2015200065

【提出日】 平成10年 6月 4日

【あて先】 特許庁長官殿

【国際特許分類】 G10L 00/00

【発明の名称】 発話文処理装置

【請求項の数】 5

【発明者】

 【住所又は居所】 大阪府門真市大字門真 1 0 0 6 番地 松下電器産業株式
会社内

 【氏名】 脇田 由実

【特許出願人】

 【識別番号】 000005821

 【氏名又は名称】 松下電器産業株式会社

【代理人】

 【識別番号】 100092794

 【弁理士】

 【氏名又は名称】 松田 正道

 【電話番号】 06 397-2840

【手数料の表示】

 【予納台帳番号】 009896

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

 【物件名】 要約書 1

 【包括委任状番号】 9006027

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 発話文処理装置

【特許請求の範囲】

【請求項1】 音声で入力される言語変換の対象となる文と、言語変換された音声またはテキストの文とが対になった学習用データベースから単語または単語列に対する文法的または意味的制約規則を学習して得られた言語規則を格納する格納手段と、

格納された前記言語規則を用いて入力音声の音声認識を行い、言語変換の対象となる文で認識結果を出力する音声認識部と、

前記音声認識部で用いられたのと同じ前記言語規則を用いて言語変換の対象となる文を言語変換された文に変換する言語変換部と、を備えたことを特徴とする発話文処理装置。

【請求項2】 前記言語規則は、言語変換の対象となる文と、変換された文とが共に意味的なまとまりを形成する部分（体型非依存フレーズと呼ぶ）に分割し、前記体型非依存フレーズ内の言語規則と前記体型非依存フレーズ間の言語規則とを分けて規則化されて作られるものであることを特徴とする請求項1記載の発話文処理装置。

【請求項3】 前記言語規則は、前記体型非依存フレーズ内の文法的または意味的規則と前記体型非依存フレーズ間の共起または接続関係を規則化されて作られるものであることを特徴とする請求項2記載の発話文処理装置。

【請求項4】 言語変換された文が同じである言語規則を同じカテゴリーとしてまとめられた言語規則群に対して、前記言語規則群に含まれる言語規則の言語変換の対象となる文の音響的規則間距離を算出する規則間距離算出部と、

音声認識の認識レベルを上げるために、算出された前記距離が近い言語規則どうしをマージすることで前記規則群の最適化を行う最適規則作成部と、を備えたことを特徴とする発話文処理装置。

【請求項5】 請求項1～4における各手段の全部または一部の機能を実現するためのプログラムを格納していることを特徴とする媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、入力音声認識し、認識結果を他言語または他の文体型などに変換して出力する発声文処理装置に関するものである。

【0002】

【従来の技術】

以下、従来の技術を発声文処理装置の1つである、入力音声了他言語に翻訳（以下通訳と呼ぶ）する装置を例にして説明する。以下通訳装置における言語変換の対象となる文を原言語文、言語変換された文を目的言語文と呼ぶ。

通訳装置は、音響信号として入力された発声文を単語テキスト列で表示された出力文に変換するための音声認識と、単語テキスト列で表示された文を入力し他言語文に翻訳する言語翻訳とを順次実行することで通訳を実現している。さらに上記言語翻訳部は、入力文の文法的または意味的構造を解析する言語解析部と、解析結果に基づいて他言語に変換する言語変換部と、翻訳結果から自然な出力文を生成する出力文生成部とから構成されている。

【0003】

しかし、音声認識部が発声文の一部を誤認識した場合や、文にあいづちや言い直しなどが挿入されたり、または文として不完結なまま発声を終えてしまうなど、発声文自体が文法的または意味的にも不自然な場合は、音声認識結果を言語解析部に入力しても解析が失敗し、結果的に翻訳結果が出力されないという問題があった。

【0004】

この問題を解決するために、発声文の一部が不自然であっても、解析が失敗しないようにする方法が提案されている。文の意味のあるまとまった部分文（以後フレーズと呼ぶ）に分割し、フレーズ内とフレーズ間とを分けて規則化し、不完結な発声にはフレーズ内規則のみを用いて解析し、解析結果の出力を可能にするように構成する方法である。（たとえば 竹沢、森元：電子通信学会論文誌D-II, Vol. J79-D-II(12)）。

【0005】

以下、文献の従来例を図5の通訳装置例に沿って説明する。

【0006】

まず通訳する前に、予め対訳コーパスや対訳単語辞書を有する学習データベース1から発声文の言語規則、翻訳規則などを学習する。ここで、対訳コーパスとは、原言語文と目的言語文が対になった文単位のデータベースであり、通常音声データも含むものである。また、対訳単語辞書とは、原言語の単語に対応する目的言語の単語が記述されているものである。例えば日英翻訳の場合では「りんご」と「apple」というように単語群が対になって格納されているものである。言語規則の学習は言語規則作成部19で行われる。図6に言語規則学習例をしめす。まず、品詞タグが付与されている対訳コーパスを用いて、各文の内容語を品詞に置き換える。すなわち、図6では、内用語の品詞化36によって、「今晚、部屋の予約をしたいんですが」34を「＜普通名詞＞|＜普通名詞＞の|＜さ変名詞＞を|したいんですが」35に置き換える。さらに各文は原言語におけるフレーズ（体型依存フレーズと呼ぶ）毎に区切られる。すなわち、目的言語のフレーズは原言語のフレーズと異なっている場合があり、目的言語のフレーズを考慮しないので、体型依存フレーズと呼ぶ。たとえば、体型依存フレーズを文節と定義した場合は、文節毎に区分けされる。図6において、34の例で言えば、「今晚」、「部屋の」、「予約を」、「したいんですが」などがフレーズの例である。次に、各文節が、品詞（内容語）と単語（機能語）との混合列とで表され、品詞化されている部分の単語名とともに、言語依存フレーズ内規則20として記述される。図6において、フレーズ内規則の記述37が、図6の20のように記述される。図6の20において、「規則番号」は、言語規則につけられた番号であり、規則は、「＜普通名詞＞」、「＜普通名詞＞の」、「＜さ変名詞＞を」、「したいんですが」等がある。その意味は文中で意味を持つ単位と考えることが出来る。例えば、規則「＜普通名詞＞」の品詞の内容は、「今晚」などがあり、他にも、「昨夜」、「明日」等も規則「＜普通名詞＞」である。ただし、これらは、対訳コーパスに文例としてあげられた場合のみ品詞の内容に加えられるものである。フレーズ内規則を対訳コーパスの全文に対して上記規則を記述した後、各フレーズ内規則の共起関係を言語依存フレーズ間規則21として記述する。こ

ここで共起関係とは一文にともに存在する関係を意味する。たとえば、共起関係をフレーズbi-gramとして規則化する場合は、各体型依存フレーズの隣接頻度が記述される。図6において、フレーズ間規則の記述38が上述したことをおこない、図6の21にフレーズ間規則が記述される。図6の21がフレーズbi-gramの例である。図6の21は規則番号対が例えば「(規則1)(規則2)」となっており、その出現頻度が10となっている。これは対訳コーパスから学習する過程で、規則1と規則2が文中にあらんで出現する回数が10回あったことを意味する。規則2と規則3が文中にあらんで出現する回数は図6の例では4回あったことになる。

【0007】

次に翻訳規則23を翻訳規則作成部22で学習する。翻訳規則の学習例を図7に示す。まず、先の言語規則の学習時と同様に、原言語および目的言語における内容語を品詞化する。図7においては、「今晚、部屋の予約をしたいんですが」と「I'd like to room-reservation tonight」39が対訳コーパスの中に格納されているものとする。そうすると内用語の品詞化41で、今度は、体型非依存フレーズに内容語を品詞化する。品詞化の結果、「＜普通名詞＞|＜普通名詞＞の＜サ変名詞＞|をしたいんですが」40となる。ここで注意すべきことは、言語規則作成19では、体型依存フレーズに分割していたので、つまり本従来例では日本語だけを考慮してフレーズに分割していたので、「＜普通名詞＞の|＜サ変名詞を」と分割されていた。ところが、翻訳規則作成部22では、体型非依存フレーズに分割しているので、つまり本従来例では日本語と英語とともに考慮して分割しているので、「＜普通名詞＞の＜サ変名詞＞を」と分割されている点である。次に、先の言語規則を学習した際の原言語における体型依存フレーズと目的言語における体型依存フレーズとが一まとまりとして対応している場合に、その一まとまりを体型非依存フレーズの境界とする。原言語の体型依存フレーズに対応する目的言語の体型依存フレーズが1まとまりとして対応しない場合には、対応する部分が一まとまりとして存在するまで体型依存フレームの連結やフレーム境界の修正を行い体型非依存フレーズとする。このようにして決められた体型非依存フレーム内の品詞（内容語）と単語（機能語）との混合列を原言語と目的言

語とを対応させながら言語非依存フレーズ内翻訳規則 23 として記述する。さらに体型非依存フレームが一文よりも短い場合には、体型非依存フレーズ間の構文構造も原言語と目的言語を対応させながら記述し、言語非依存フレーズ間翻訳規則 25 に追加する。図 7 においては、フレーズ内規則の記述 42 により 23 のごとく記述される。規則 2 は、規則が日本語では「＜普通名詞＞の | ＜サ変名詞を」である。また、英語では、「＜noun＞」である。その対訳コーパスにおける出現頻度は 7 回で、品詞の内容としては、日本語では「(部屋)、(予約)」となり、英語では、「(room reservation)」となる。またフレーズ間規則の記述 43 が、日本語と英語で品詞が現れる順序が違っているので、順序関係の対応をつけるために 25 で言語構造をツリー状にして対応をとっている。

【0008】

文生成規則 5 には、コーパスから作成された翻訳規則で不足している一般的な目的言語規則を記述しておく。たとえば、日英翻訳の場合には、冠詞および不定冠詞規則や三人称単数化規則などがその内容として記述されている。

通訳の際には、まず発声された原言語音声はマイクロホン 6 から入力され、音声認識部 7 に入力される。音声認識部 7 では、たとえば、言語依存フレーズ内言語規則 20 として記述されている品詞と単語の混合列および言語依存フレーズ間言語規則 21 としてのフレーズの隣接頻度とにより、時系列に沿って順次認識単語候補が予測される。予め学習されている音響モデル 8 と入力音声との距離値をベースとした音響スコアとフレーズ bi-gram による言語スコアとの和を認識スコアとし、Nbest-search により認識候補である連続単語列が決定される。

【0009】

認識された連続単語列は、言語翻訳部 24 に入力される。言語翻訳部 24 では、入力された連続単語列から翻訳規則内の体型非依存フレームを探索し、対応する目的言語の構文構造または品詞と単語の混合列と対訳単語辞書を用いて、入力された原言語の連続単語列を目的言語文に変換する。

【0010】

変換された目的言語文は出力文生成部 10 に入力され、文法的な不自然さを修正する。たとえば、定冠詞や不定冠詞の付与、代名詞、動詞における 3 人称化や

複数化や過去形化などの最適化などが行われる。修正後の目的言語翻訳結果文はたとえばテキストとして出力される。

【0011】

【発明が解決しようとする課題】

しかしながら、従来例における会話文処理装置である通訳装置は、体型依存フレーズの単位と体型非依存フレーズの単位とが異なっている場合が多く、原則的には体型依存フレーズが連結されて体型非依存フレーズとなっていて、体型非依存フレーズの方が単位が長くなる場合が多い。従って、学習された規則が当てはまらない部分を持つ発声文が入力されたり、文の一部が誤認識した場合には、言語規則として受理できた音声認識結果フレーズ列が言語翻訳部では受理できない場合がある。そのため、フレーズ毎に解析された結果が、言語変換で失敗し、翻訳結果が出力されないという問題が生じる。

【0012】

本発明はこのような従来の発話文処理装置の一部の未学習部分または誤認識部分があるために、文全体の言語変換が失敗し結果が全く出力されないという課題を考慮し、入力音声文に未学習部分があったり、音声認識が一部誤りを起こしても、必ず目的体型への変換を可能とし、従って正しく認識され適切な解析規則が当てはまった部分については、部分的にでも正しい結果を出力できる会話文処理装置を提供することを目的とするものである。

【0013】

【課題を解決するための手段】

上述した課題を解決するために、第1の本発明（請求項1に対応）は、音声で入力される言語変換の対象となる文と、言語変換された音声またはテキストの文とが対になった学習用データベースから単語または単語列に対する文法的または意味的制約規則を学習して得られた言語規則を格納する格納手段と、

格納された前記言語規則を用いて入力音声の音声認識を行い、言語変換の対象となる文で認識結果を出力する音声認識部と、

前記音声認識部で用いられたのと同じ前記言語規則を用いて言語変換の対象となる文を言語変換された文に変換する言語変換部と、を備えたことを特徴とする

発話文処理装置である。

【0014】

また第2の本発明（請求項2に対応）は、前記言語規則は、言語変換の対象となる文と、変換された文とが共に意味的なまとまりを形成する部分（体型非依存フレーズと呼ぶ）に分割し、前記体型非依存フレーズ内の言語規則と前記体型非依存フレーズ間の言語規則とを分けて規則化されて作られるものであることを特徴とする第1の発明に記載の発話文処理装置である。

【0015】

また第3の本発明（請求項3に対応）は、前記言語規則は、前記体型非依存フレーズ内の文法的または意味的規則と前記体型非依存フレーズ間の共起または連接関係を規則化されて作られるものであることを特徴とする第2の発明に記載の発話文処理装置である。

【0016】

また第4の本発明（請求項4に対応）は、言語変換された文が同じである言語規則を同じカテゴリーとしてまとめられた言語規則群に対して、前記言語規則群に含まれる言語規則の言語変換の対象となる文の音響的規則間距離を算出する規則間距離算出部と、

音声認識の認識レベルを上げるために、算出された前記距離が近い言語規則どうしをマージすることで前記規則群の最適化を行う最適規則作成部と、を備えたことを特徴とする発話文処理装置である。

【0017】

また第5の本発明（請求項5に対応）は、第1～4の発明における各手段の全部または一部の機能を実現するためのプログラムを格納していることを特徴とする媒体である。

【0018】

次に本発明の動作を説明する。

【0019】

第1の本発明記載の発話文処理装置においては、入力される音声の言語変換の対象となる文と言語変換された文とが対になった学習用データベースから、言語

変換の対象と言語変換された対象とが対応した形で言語規則を記述しておき、この言語規則を用いて入力音声の連続単語認識を行うことにより、必ず言語変換された文に変換可能な認識結果を出力でき、従って、入力文の一部が未知部分文であったり、音声認識が一部誤ったとしても、正しく認識および解析された部分は適切に処理され出力されることを可能にする。

【0020】

また、第2の本発明記載の発話文処理装置における言語規則は、入力文を体型非依存フレーズに分割し、体型非依存フレーズ内の言語規則と体型非依存フレーズ間の言語規則とを分けて規則化されて作られるものであり、入力音声文に未学習部分があったり、音声認識が一部誤りを起こしても、正しく認識され適切な解析規則が当てはまった部分のみの変換が可能であり、部分的な変換結果を必ず出力することを可能にする。

【0021】

また、第3の本発明記載の発話文処理装置における言語規則は、体型非依存フレーズ内の意味的規則と体型非依存フレーズ間の共起または接続関係を規則化されて作られるものであり、発話文が文法的に曖昧であっても、正しく認識することが可能となる。

【0022】

また、第4の本発明記載の発話文処理装置においては、解析された言語規則の音響的な規則間距離を算出する規則間距離算出部と、算出された規則間距離に近い規則をマージすることで規則の最適化を行うことにより、言語規則の性能をなるべく落とさず、体型非依存フレーズを用いることの規則数の増加を押さえることが可能となる。従って、高性能な音声認識及び目的体型変換の実現が可能となる。

【0023】

【発明の実施の形態】

以下に本発明の実施の形態について図面を参照して説明する。

【0024】

まず第1の実施の形態について説明する。

【0025】

第1の実施の形態では、発話文処理装置の一例として、従来例同様、異なる言語間の変換を行う通訳装置を用いて説明する。図1は本発明の請求項1、請求項2、請求項3に係る一実施の形態である通訳装置のブロック図である。

【0026】

本実施の形態の通訳装置は、まず通訳する前に、言語解析部2で予め対訳コーパスや対訳単語辞書などを有している学習用データベース1から発声文の原言語及び目的言語の言語規則を学習する。言語規則の学習例を図3に示す。言語規則作成部2では、たとえば、品詞タグが付与されている対訳コーパスを用いて原言語文及び目的言語文の内容語を品詞化する。図3において、対訳コーパスの文「今晚、部屋の予約をしたいんですが」「I'd like to room-reservation tonight」26が、内容語の品詞化30で、「＜普通名詞＞|＜普通名詞＞の＜サ変名詞＞|をしたいんですが」27のように品詞化される。さらに、原言語におけるフレーズと目的言語におけるフレーズとが一まとまりとして対応している場合に、その一まとまりを体型非依存フレーズとしてその境界を区切る。これは従来例の翻訳規則の学習と同様である。図3においては、27の「＜普通名詞＞の＜サ変名詞＞」が上述した処理の結果である。従来例との大きな違いは次のような点である。すなわち従来例では、言語規則の作成19で体型依存フレーズを用いて内用語の品詞化を行っていたが、本実施の形態では、最初から言語規則の作成2において体型非依存フレーズを用いて内用語の品詞化を行う点である。次に各体型非依存フレーズにおいて、品詞と単語の混合列、および品詞で表されている部分の単語名、さらに各体型非依存フレーズの対訳コーパスにおける出現頻度を体型非依存フレーズ内規則3として記述する。対訳コーパスの全文に対して上記規則を記述する。図3においては、上述した内容は、フレーズ内規則の記述31により3に記述される。図3の3において、規則1は、日本語が「＜普通名詞＞」であり、英語が「＜noun＞」である。品詞の内容としては、日本語が「今晚」、英語が「tonight」となっている。対訳コーパスに現れていれば、「明日」、「tomorrow」等も規則1に記述されるものである。各フレーズ内規則の共起関係を体型非依存フレーズ間規則4として記述する。たとえば、共起関係をフレーズbi-g

ramとして規則化する場合は、各体型非依存フレーズの隣接頻度を記述しておく。上述した内容は、図3において、フレーズ間規則の記述32が、28を記述することを意味する。28がフレーズbi-gramのことであり、その意味は従来例の翻訳規則の学習を記述した部分と同一である。さらに、各体型非依存フレーズ間の構文構造も体型非依存フレーズ間規則4に記述しておく。これは図3において、フレーズ間規則の記述32が29を記述することである。その意味は従来例の翻訳規則の学習を記述した部分と同一である。

【0027】

文生成規則5には、上記言語規則3および4で不足している目的言語規則を記述しておく。たとえば、日英翻訳の場合には、冠詞および不定冠詞規則や三人称単数化規則などがその内容として記述されている。これは従来例の文生成規則と同様に動作するものである。

【0028】

なお、フレーズ内言語規則3及び／またはフレーズ間言語規則4が本発明の格納手段の例である。

【0029】

通訳の際には、まず発声された原言語音声はマイクロホン6から入力され音声認識部7に入力される。音声認識部では、たとえば、体型非依存フレーズ内言語規則3として記述されている品詞および列単語の混合列と体型非依存フレーズ間言語規則4としてのフレーズbi-gramとにより、時系列に沿って順次認識単語候補が予測される。予め学習されている音響モデル8と入力音声との距離値をベースとした音響スコアとフレーズbi-gramによる言語スコアとの和を認識スコアとし、Nbest-searchにより認識候補である連続単語列が決定される。このように決定された連続単語列は言語変換部9に入力される。フレーズ内言語規則3、フレーズ間言語規則4では、予め原言語と目的言語とが対応しながら規則化されている。言語変換部9では、上記規則を用いて、本連続単語列は目的言語のフレーズ列に変換され出力される。この際、入力された原言語フレーズ列が、既に学習されたフレーズ間の構文構造に当てはまる場合には、目的語のフレーズ列は構文構造に沿って修正された後出力される。

出力された目的言語文は出力文生成 10 に入力され、文法的な不自然さを修正する。この動作は従来例の出力文生成 10 と同様である。

【0030】

以上の実施の形態では、音声認識で使用する言語規則を学習する際に、原言語と目的言語とがともに意味をもつ一かたまりとなった部分を単位として規則化を行い、この規則の制約に基づいて認識を行うことにより、入力音声文に未学習部分があったり、音声認識が一部誤りを起こしても、全文に対する翻訳結果が全く出力されないという問題点を解決し、正しく認識された部分については、適切な翻訳結果を出力できる会話文処理装置を実現できる。

【0031】

なお、本実施の形態では、会話文処理装置の 1 つの例として通訳装置を例にあげて説明したが、これは他の会話文処理装置、例えばくだけた発話文を書き言葉のようなテキスト文に変換する会話文処理装置においても、発話文を原体型、テキスト文を目的体型として、同様に使用できるものである。

【0032】

次に第 2 の実施の形態について図面を参照しながら説明する。本実施の形態でも、第 1 の実施の形態同様、通訳装置を用いて説明する。図 2 は本発明の請求項 4 に係る一実施の形態である通訳装置のブロック図である。

本実施の形態の通訳装置は、まず通訳する前に、予め言語規則作成部 11 で対訳コーパスや対訳単語辞書を有している学習データベース 1 から発声文の原言語及び目的言語のフレーズ内言語規則 12、フレーズ間言語規則 13 を学習する。学習される規則は、第 1 の実施の形態における言語規則の学習と同様である。次に学習された言語規則の最適化を行う。最適化の例を図 4 に示す。

【0033】

まず、学習された体型非依存フレーズにおいて、目的言語フレーズが同じであるフレーズを同カテゴリーとしてまとめる。図 4 において、12 は言語規則であり、規則間距離算出 14 で、33 のごとくカテゴリーとしてまとめる。規則 1、規則 2 規則 3 は目的言語規則が「I'd like to」と同じであるので、同カテゴリーになる。また、規則 4 は、目的言語規則が「please」となっているので、規則

1、規則 2、規則 3 とは別のカテゴリーに分類される。次に同カテゴリーに含まれる原言語フレーズ間の音響的距離を規則間距離算出部 14 で算出する。図 4 において、15 が原言語フレーズ間の音響的距離を算出した例である。15 では、規則 1 と規則 2 の距離は 7 となっており、規則 1 と規則 3 の距離は 2 となっている。

【0034】

同カテゴリー規則における原言語フレーズの音響的距離は次のように算出する。まず、カテゴリー内の全ての目的言語フレーズにおける混合列の品詞部分に、同品詞であれば同じ単語を当てはめ、全ての混合列を単語列に変換する。次に各単語列の発音が類似しているかを調べるために、各単語列の文字列の違いに対する距離を、数 1 を用いて算出し、規則間距離テーブル 15 に記述する。n 個の単語からなるフレーズ $X = \{x_1, x_2, x_3, \dots, x_n\}$ (x は各単語) と m 個の単語からなるフレーズ $Y = \{y_1, y_2, y_3, \dots, y_m\}$ との間の距離を $D(X_n, Y_m)$ として、

【0035】

$$\begin{aligned} \text{【数 1】 } D(X_i, Y_j) = & \min |D(X_{i-1}, Y_j) + d(x_i, y_j)| \\ & |D(X_{i-1}, Y_{j-1}) + d(x_i, y_j)| \\ & |D(X_i, Y_{j-1}) + d(x_i, y_j)| \end{aligned}$$

但し、if $x_i = y_j$ then $d(x_i, y_j) = 0$
 else $d(x_i, y_j) = 1$

次に最適規則作成部 16 で、距離値が一定値以内であるフレーズの中で、最も出現数の多い規則のみを残し、他の規則を消去する。たとえば、図 4 の例では、上記一定値を 2 とした場合、33 において、同カテゴリーである規則 1 と規則 3 との規則間距離は 2 であり、上記一定値 2 以下である。従って、この 2 つの規則の出現頻度の多い規則 1 を採用し、規則 3 を規則から削除する。それに合わせて出現数も書き換える。

【0036】

フレーズ内言語規則 12 に書かれている全ての規則に対して上記最適規則化を行った後、消去されなかった言語規則のみをフレーズ内最適言語規則 17 として保管する。最適化された規則に従い、フレーズ間規則 13 の中の除去された規則

を採用した規則で書き換え、合わせて出現数も修正する。図4において、最適規則作成16により規則3は削除され、規則1として1本化される。それにあわせて、規則1の出現数は17のごとく削除された規則3との和である15となっている。

【0037】

文生成規則5には、コーパスから作成された上記言語規則で不足している目的言語規則を記述しておく。たとえば、日英翻訳の場合には、冠詞および不定冠詞規則や三人称単数化規則などがその内容として記述されている。これは従来例の文生成規則と同様に動作するものである。

【0038】

通訳の際には、まず発声された原言語音声はマイクロホン6から入力され音声認識部7に入力される。音声認識部では、たとえば、体型非依存フレーズ内言語規則17として記述されている品詞および列単語の混合列と体型非依存フレーズ間言語規則18としてのフレーズ隣接頻度とにより、時系列に沿って順次認識単語候補が予測される。予め学習されている音響モデル8と入力音声との距離値をベースとした音響スコアとフレーズbi-gramによる言語スコアとの和を認識スコアとし、Nbest-searchにより認識候補である連続単語列が決定される。このように決定された連続単語列は言語変換部9に入力される。言語規則17、18では、予め原言語と目的言語とが対応しながら規則化されている。言語変換部9では、上記規則を用いて、本連続単語列は目的言語のフレーズ列に変換され出力される。この際、入力された原言語フレーズ列が、既に学習されたフレーズ間の構文構造に当てはまる場合には、目的語のフレーズ列は構文構造に沿って修正された後出力される。

【0039】

出力された目的言語文は出力文生成部10に入力され、文法的な不自然さを修正する。この動作は従来例の出力文生成部10と同様である。

以上の実施の形態では、音声認識で使用する言語規則を学習する際に、原言語と目的言語とがともに意味をもつ一かたまりとなった部分を単位として規則化を行った後、規則化されている目的言語部分が同じである原言語フレーズが音響的に

類似している場合には、類似している中から最も出現頻度の高い規則のみを採用し残りの規則を消去することにより、なるべく言語規則の性能を落とさずに、体型非依存フレームを単位にすることによる規則数の増加を押さえ、従って高性能な認識及び言語変換を可能にする通訳装置を実現するものである。

【0040】

なお、本実施の形態では、会話文処理装置の1つの例として通訳装置を例にあげて説明したが、これは他の会話文処理装置、例えば発話文をテキスト文に変換する会話文処理装置においても、発話文を原体型、テキスト文を目的体型として、同様に使用できるものである。

【0041】

なお、本発明は、その機能を実現する各手段の全部または一部の機能を実現するためのプログラムを格納していることを特徴とする媒体でもある。

【0042】

【発明の効果】

以上説明したところから明らかなように、請求項1～3の本発明は、必ず処理可能な認識結果を出力することが可能となり、従って、発声文に一部の未学習部分があったり、認識結果に一部誤認識部分があっても、文全体の言語変換が失敗し結果が全く出力されないという問題がなく、正しく認識された部分は適切に目的体型に変換することが可能になる発話文処理装置を提供することができる。

【0043】

さらに、請求項2の本発明は、入力音声文に未学習部分があったり、音声認識が一部誤りを起こしても、正しく認識され適切な解析規則が当てはまった部分のみの変換が可能であり、部分的な変換結果を必ず出力することを可能にする発話文処理装置を提供することができる。

【0044】

さらに、請求項3の本発明は、発話文が文法的に曖昧であっても、正しく処理することが可能となる発話文処理装置を提供することができる。

【0045】

さらに、請求項4の本発明は、言語規則の性能をなるべく落とさず、体型非依

存フレーズを用いることによる言語規則数の増加を押さえることが可能となり、高性能な音声認識及び目的体型変換の実現が可能となる発話文処理装置を提供することができる。

【図面の簡単な説明】

【図 1】

本発明の第 1 および第 2 の実施の形態における構成図。

【図 2】

本発明の第 2 の実施の形態における構成図。

【図 3】

本発明の第 1 および第 2 の実施の形態における体型非依存フレーズ内言語規則と体型非依存フレーズ間言語規則の作成例を示す図。

【図 4】

本発明の第 2 の実施の形態における最適規則の作成例を示す図。

【図 5】

従来例の構成図。

【図 6】

従来例の音声認識のための体型依存フレーズ内言語規則と体型依存フレーズ間規則の作成例を示す図。

【図 7】

従来例の言語翻訳のための体型非依存フレーズ内言語規則と体型非依存フレーズ間規則の作成例を示す図。

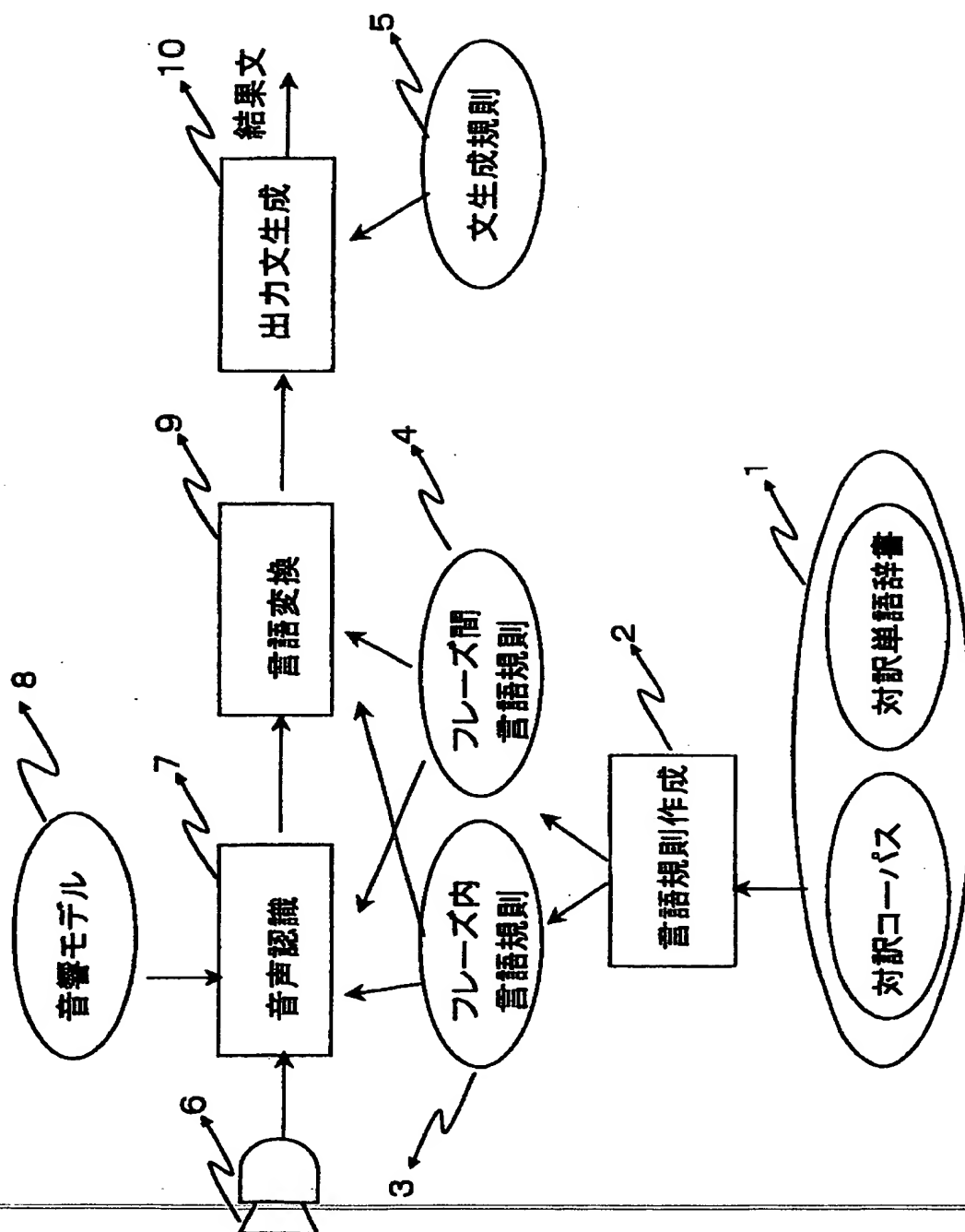
【符号の説明】

- 1 学習データベース
- 2、 1 1 言語解析部
- 3、 1 2 フレーズ内言語規則
- 4、 1 3 フレーズ間言語規則
- 5 文生成規則
- 6 マイクロホン
- 7 音声認識部

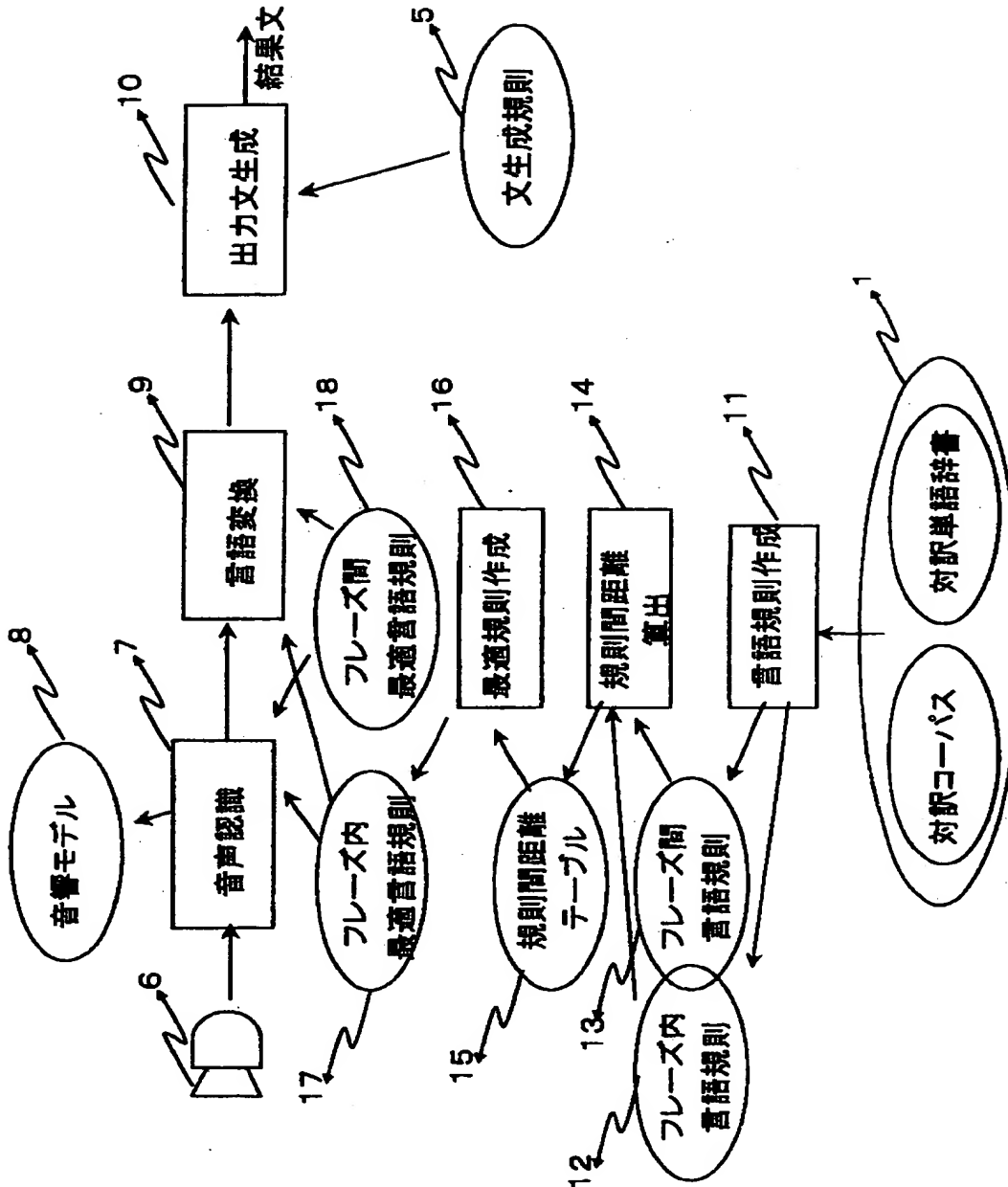
- 8 音響モデル
- 9 言語変換部
- 10 出力文生成
- 14 規則間距離算出部
- 15 規則間距離テーブル
- 16 最適規則作成
- 17 フレーズ内最適言語規則
- 18 フレーズ間最適言語規則

【書類名】 図面

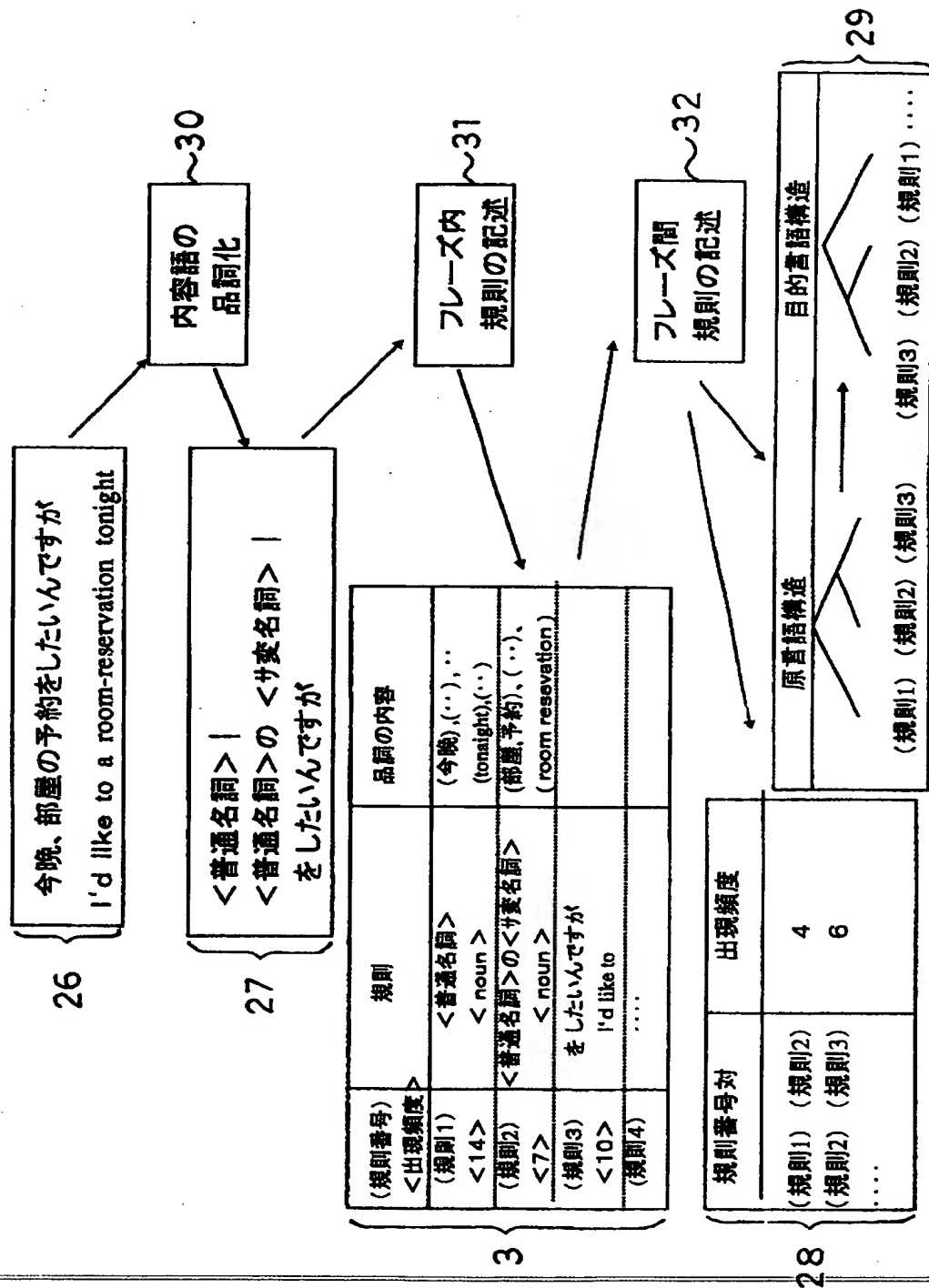
【図 1】



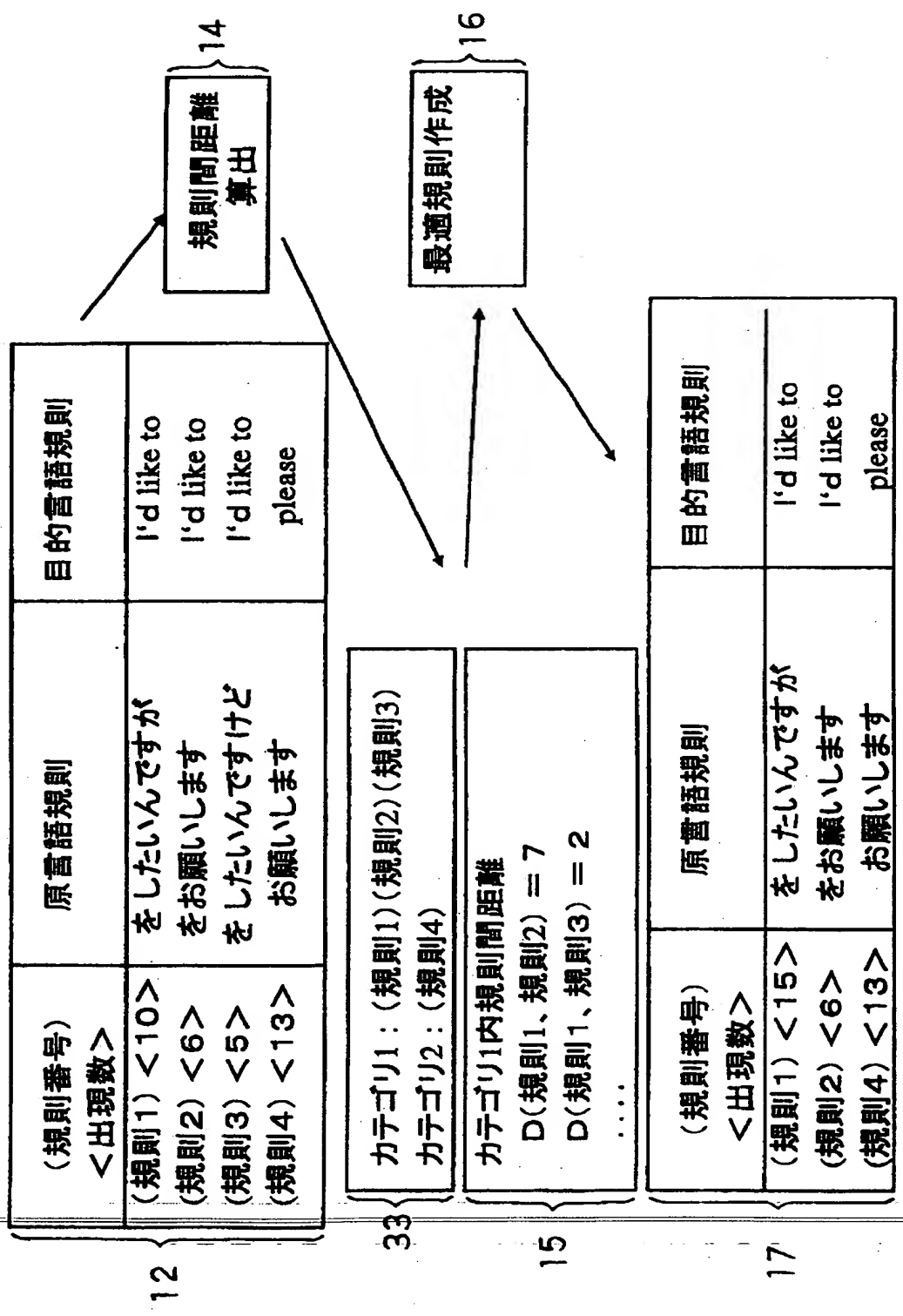
【図 2】



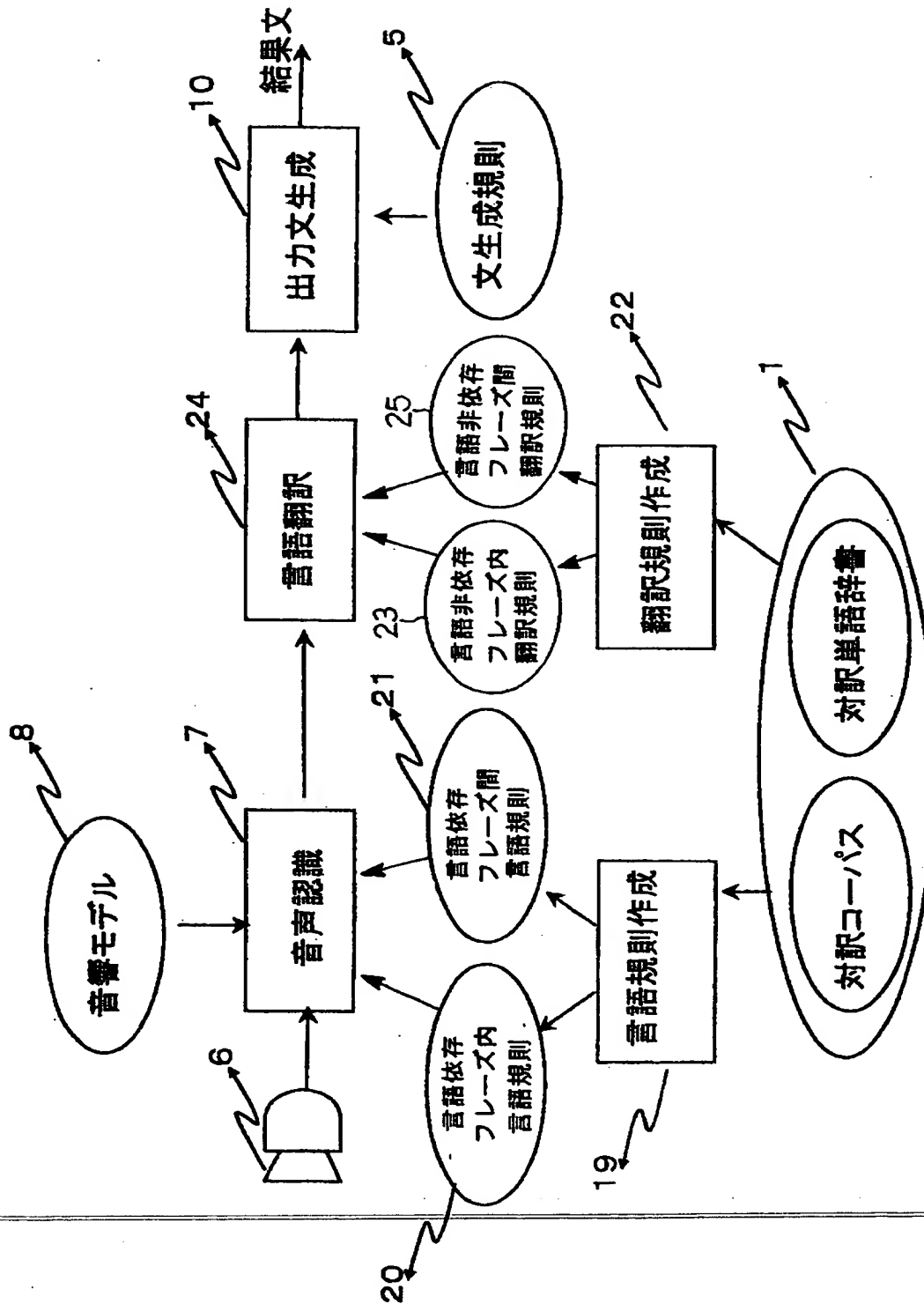
【図 3】



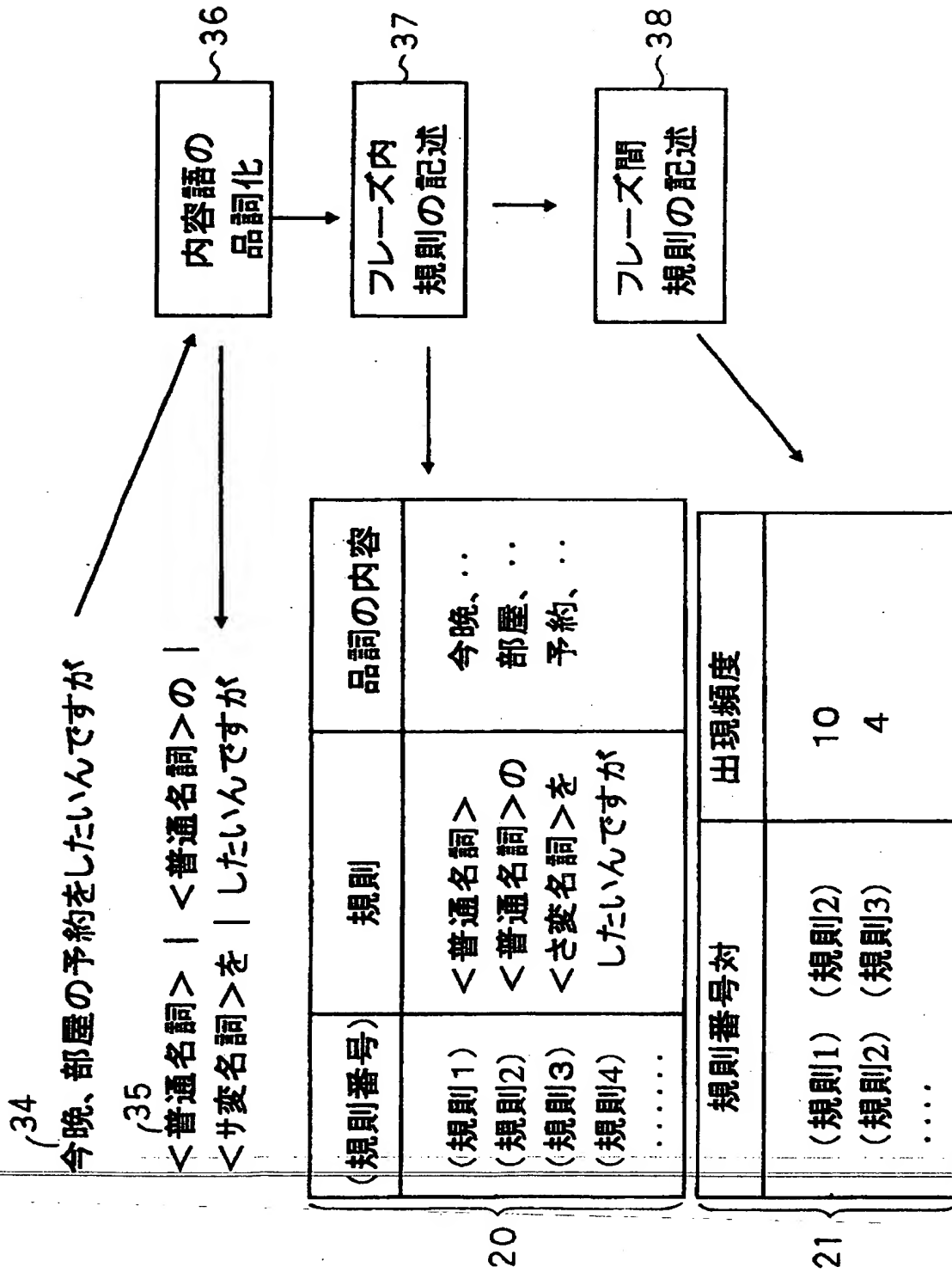
【図4】



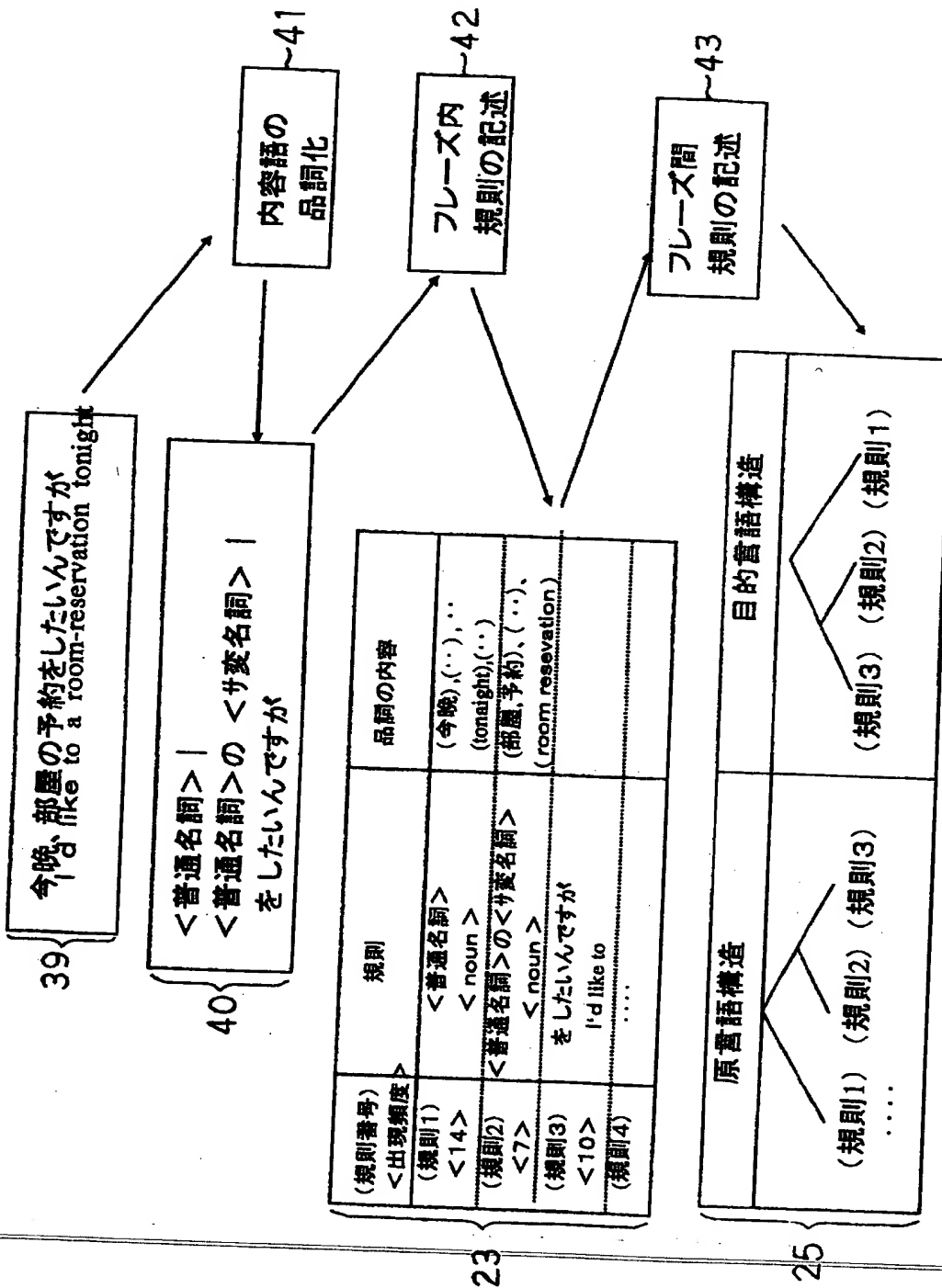
【図 5】



【図6】



【図 7】



【書類名】 要約書

【要約】

【課題】 入力された音声を音声認識し、その結果を別の言語に翻訳する発話文処理装置において、一部の未学習部分または誤認識部分があるために、文全体の翻訳が失敗し結果が全く出力されないという課題がある。

【解決手段】 音声で入力される言語変換の対象となる文と、言語変換された音声またはテキストの文とが対になった学習用データベースから単語または単語列に対する文法的または意味的制約規則を学習して得られた言語規則を格納する格納手段 3 または 4 と、格納された前記言語規則を用いて入力音声の音声認識を行い、言語変換の対象となる文で認識結果を出力する音声認識部 7 と、前記音声認識部で用いられたのと同じ前記言語規則を用いて言語変換の対象となる文を言語変換された文に変換する言語変換部 9 と、を備えたことを特徴とする発話文処理装置である。

【選択図】 図 1

【書類名】 職権訂正データ
【訂正書類】 特許願

<認定情報・付加情報>

【特許出願人】

【識別番号】 000005821

【住所又は居所】 大阪府門真市大字門真 1006 番地

【氏名又は名称】 松下電器産業株式会社

【代理人】 申請人

【識別番号】 100092794

【住所又は居所】 大阪市淀川区宮原 5 丁目 1 番 3 号 新大阪生島ビル

松田特許事務所

【氏名又は名称】 松田 正道

出 願 人 履 歴 情 報

識別番号

[000005821]

1. 変更年月日

1990年 8月28日

[変更理由]

新規登録

住 所

大阪府門真市大字門真1006番地

氏 名

松下電器産業株式会社